JAN 4 195

ol 10 No 2 (New Series No 14) December 1949

ANALYSIS

Edited by
Margaret Macdonald

B
PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

with the advice of

A. J. Ayer A. E. Duncan-Jones
R. B. Braithwaite C. A. Mace
Herbert Dingle A. M. MacIver
H. H. Price

CONTENTS

TWO SHILLINGS NET

Annual subscription 10s. 6d. post free for six numbers

BASIL BLACKWELL · BROAD STREET · OXFORD

OL
O
S
E
49
MI

ANALYTIC—SYNTHETIC

I

By FRIEDRICH WAISMANN

1. What is " analytic " ?

KANT says[1] " In all judgments in which there is a relation between subject and predicate . . . that relation can be of two kinds. Either the predicate B belongs to the subject A as something contained (though covertly) in the concept A ; or B lies outside the sphere of the concept A, though somehow connected with it. In the former case I call the judgment analytical, in the latter synthetical. Analytical judgments (affirmative) are therefore those in which the connection of the predicate with the subject is conceived through identity, while others in which that connection is conceived without identity, may be called synthetical. The former might be called illustrating, the latter expanding judgments, because in the former nothing is added by the predicate to the concept of the subject, but the concept is only divided into its constituent concepts which were always conceived as existing within it, though confusedly ; while the latter add to the concept of the subject a predicate not conceived as existing within it, and not to be extracted from it by any process of mere analysis . . . It is clear from this that our knowledge is in no way extended by analytical judgments, but that all they effect is to put the concepts which we possess into better order and render them more intelligible ".

This definition may seem clear enough ; yet isn't it surprising how easy it is to raise questions which are plainly embarrassing ? What, for instance, is meant by saying that, in an analytic judgment, the concept of the subject is " only *divided into* its constituent concepts " ? Is the subject-term to be regarded as a sort of sum total of its constituent concepts, i.e., of all those which, analytically, can be asserted of it ? This doesn't seem to make sense. Suppose we make a number of analytic judgments such as S is P_1, S is P_2, . . . S is P_n, these being *all* the analytic judgments we can make about this particular subject S. Then, according to Kant, it would seem natural to say that S is thereby " divided ", or " dissolved "[2], into the constituent concepts P_1, P_2, . . . P_n But what meaning are we to attach to this ? Shall we say, e.g.,

[1] Critique of Pure Reason : Introduction IV.
[2] In German " *durch Zergliederung zerfallt* " (Critique), " *aufgelost* " (Proleg.).

that S is the collection, or the class, or the totality of all the members $P_1, P_2, \ldots P_n$? To assert of a class any of its member concepts, e.g., to assert of the class $\{P_1, P_2, \ldots P_n\}$ that it is P_1, would be absurd. We don't say of the attributes of a thing that *they*, all of them, have any of these attributes. Take the simplest case in which the class consists of one member only, say P. Then the class consisting of P will be different from P, and to say of this class $\{P\}$ that it is itself P, i.e. that it has the property expressed by " P ", would be self-stultifying, or more accurately, it would be a type-fallacy, landing us in a lot of well-known logical contradictions. So this way is blocked : the relation of S to any of its predicates must be different from the relation of a class to any of its members. In what, then, *does* it consist ? The same difficulty can be brought out in a slightly different way when we ask—What is meant by saying that the predicate is " contained in " the concept of the subject, or is " conceived as existing within it ", or that the former can be " extracted from it "—turns of speech which sound as if they were taken from dentistry ? The word " contain ", even if not taken in its strictly spatial sense " to enclose ", is used in a great many different ways as when I say, " A pound contains 16 ounces ", " This book contains some valuable information ", " The premiss of an inference contains the conclusion ", and the like. In *which* of these senses, then, is the predicate to " be contained in " the notion of the subject ? Perhaps in the sense of our last example, that is, in the sense in which a conclusion is often held to " be contained in " the premiss from which, according to the same view, it can be " extracted " by inference ? That, it seems, is an analogy worth following. But if so, what exactly *is* the relation that holds between the premiss and the conclusion of an inference ? Shall I say, e.g., that whenever I think of the premiss, I *coincidently* think of the conclusion ? That would be glaringly untrue : I may consider the premiss *without* noticing some particular conclusion. (How on earth could a mathematical discovery be thrilling ?) Saying that the latter is " contained in " the former, whatever this may mean, can certainly not be taken to refer to any *psychological* relationship between the two such that thinking of the one is accompanied, or followed, by thinking of the other. No ; in studying the logical relation between the parts of an inference, we are clearly *not* investigating what is actually going on, or what may be going on, in someone's mind. Similarly in our case : to say, in the case of an analytic judgment, that the predicate is " contained in " the concept of the subject can certainly not be taken

to mean that whoever thinks of the subject will simultaneously, or a little while later, think of the predicate, or something of this sort. So far our result is mainly negative: we see that the relation can certainly not be a *psychological* one, though we are baffled to say *what* it is. I don't think that these questions are raised wantonly; they present themselves to anyone who tries to understand clearly what an analytic judgment is. Our difficulty is, at least in part, due to the fact that Kant, when he speaks of "analytic", is unwittingly using nothing but metaphorical terms, which hint at, and at the same time obscure, what the true relation is. Nor is this the only problem with which we are confronted. Another difficulty arises from the fact that Kant's definition is such as to apply, according to his own words, only to statements of the subject-predicate pattern, ignoring other types—such as relational, and existential statements, let alone mathematical formulae like $7+5=12$ (which Kant, oddly enough, cites as an example of a synthetic judgment). Subject and predicate are, after all, ideas which are borrowed from the grammar of certain word sentences, but which cannot, without destroying the clarity of our ideas, be applied to forms so fundamentally different. In other words, Kant's definition is *too narrow*, a fact we shall do well to keep in mind too.

Attempts have been made to amend Kant's definition. To quote a recent writer on the subject, A. Pap[3]: "Analytic statements . . . may be roughly characterized as statements whose truth *follows* (my italics) from the very meaning of their terms". Here I immediately come up against a stumbling block: what can be meant by saying that a statement *follows from the very meaning of its terms?* I should have thought that *one* statement can follow from *another;* but from the meaning—! Yet, strangely enough, such a view has been taken by no one less than Frege. Criticising the formalist account of mathematics—the view according to which mathematics is but a sort of game played with inkmarks on paper instead of with chessmen on a board—he very emphatically says: "If there were any meaning to be considered, the rules [of this game] could not be arbitrarily laid down. On the contrary, the rules follow necessarily from the meaning of the marks".[4] Frege's idea seems, roughly, to have been this: If I write down a rule, e.g., the equation "$2+2=4$", this may be regarded as a configuration in a game not so very unlike a configuration of chessmen on a board. But then the question of truth and falsehood does not arise. As soon as I

[3] "Indubitable Existential Statements", *Mind*, 1946.
[4] Grundgesetze der Arithmetik, Vol. ii, 1900, p. 158.

come to *know*, however, what the marks "2", "+", "=", "4" mean, I am no longer free to choose any configuration I please to include in my game: the truth of that equation seems rather to be *grounded in* the meaning of those marks. It would lead us far to dig down to the roots from which Frege's mistake springs. One word, however, will not be amiss. Whoever says that a rule, say, an equation, *follows* from the meaning of its terms, is bound to make quite clear what he understands by that. If someone tells me that the rule, "'John' is spelt with a capital" follows from, or is a consequence of, the rule, "Any proper name is spelt with a capital", I have no difficulty in understanding what he means. If, on the other hand, he tells me that an equation follows from the meaning of its terms, or that an analytic statement is one whose truth follows from the meaning of its terms, I am absolutely at a loss to make head or tail of it. One thing, at any rate is clear: as the meaning is not a starting point for making deductions, "follows", in this context, can not mean "logically follows"; so what *can* it mean? Queer that so subtle a mind as Frege should have failed to see that there is a problem, dropping not the slightest hint as to what he had in mind.

A different approach is made by M. Schlick.[5] The predicate "is contained in the concept of the subject", he says, "can only mean that it is part of its definition". This interpretation has two merits: first, it is not open to the objections we had to raise against Frege; and secondly, it can easily be extended so as to cover other types of judgments. Schlick thus arrives at the following formulation: A judgment is analytic if the ground for its truth lies solely in the *definitions* of the terms which occur in it. "Consequently, one may say with Kant that analytic judgments rest upon the law of contradiction, they derive from definitions by means of this law".[6] What Schlick obviously has in mind is that a judgment, or a statement, is analytic if it follows from mere definitions only by logical inference. Similarly Ewing (*A Short Commentary on Kant's Critique of Pure Reason,* 1938): "An analytic judgment is one which *follows* (my italics) from the definition of its subject-term."

Before proceeding to deeper lying questions let us see whether the last two definitions are satisfactory. It undeniably marks a great advance in them that they refer, not to such elusive and questionable entities as the "meaning of the terms",

[5] *Allgemeine Erkenntnislehre*, 1st ed., 1918, p. 97.
[6] *loc. cit.*

but to *definitions*. Before, however, we acquiesce in this view we are bound to ask—Is it really free from any obscurities? Strange as it may seem, when I attend to the question a bit more closely, I become doubtful. If an analytic statement is characterized as one that follows from mere definitions, why is it not itself a definition? A definition behaves in many respects like a rule, e.g., a rule of chess: it is *pre*scriptive rather than *de*scriptive—it tells us how a word, or a symbol, *is* to be used, not what its actual, or predominant, usage is. If I wish to assert that a definition given is in accord with the actual, or the prevailing, use of language, then I am, truly or falsely, making a *statement*, and no longer laying down a mere definition. There are further analogies between a definition and a rule. Thus a definition may be employed in the *learning* of a language in much the same way in which, say, the rules of chess may be employed in learning chess; a definition may be referred to in order to *justify* a certain use of words in a way similar to that in which a rule of chess may be referred to in order to justify a certain move. A definition, like a rule, can be *set up* at a certain time, and *abandoned*, or *altered*, later on; it may be *recognized*, or it may not; a definition, like a rule, can be *observed*, or be *infringed* in practice; and there are perhaps more such features. In view of this far-reaching analogy it seems very odd that it should break down at one point: what follows from a rule will, generally, be another rule (see the example concerning spelling); why, then, is it that what follows from a definition is, not, as one would expect, a *definition*, but an analytic *judgment?* Why suddenly this difference?

Let us first take a look at one or two examples in order to throw some light on the matter. Suppose I define a dragon as a fabulous winged serpent breathing flame, and a serpent, in its turn, as a scaly reptile, then I can derive from these two definitions the sentence "A dragon is a fabulous winged and scaly reptile, breathing flame", which, to all appearance, might well pass for an alternative definition of "dragon"; certainly it is not the sort of thing which would be called "analytic". Again, if I define mephitis as poisonous stench, poisonous as a property that causes harm to life, or death, and stench as a foul smell, I am led on to say that mephitis is a foul smell that causes harm to life, or death—a sentence which, in ordinary circumstances, will be taken by most people as a mere re-wording, or a more explicit form, of the definition.

So far we were considering examples culled from word language. Before jumping to conclusions, let us be cautious and

look at different sorts of examples, for instance, from symbolic logic, or from arithmetic. Suppose I start with the following three definitions :—

$$p \supset q . = . \sim p \vee q \quad \text{Df} \qquad (1)$$
$$p \vee q . = . \sim p | \sim q \quad \text{Df} \qquad (2)$$
$$\sim p . = . p | p \quad \text{Df} \qquad (3)$$

Add to these the theorem

$$\vdash . \sim \sim p . \equiv . p \qquad (4)$$

Before consequences from these formulae can be derived, certain rules of inference must first be laid down ; let the following ones be chosen :—

(I) The relation expressed by the sign " = " is symmetrical and transitive.

(II) Two equivalent expressions may be substituted for each other in any occurrence.

(III) The statement variables " p ", " q ", etc. may be replaced by other, and possibly complex, statement variables. (Rule of substitution).

Applying these three rules, we may transform $p \supset q$ as follows:

$$p \supset q . =_{(1)} . \sim p \vee q . =_{(2)} \sim \sim p | \sim q . =_{(4)} . p | \sim q . =_{(3)} p | q | q .$$

Here each step of the transformation is made according to the rules (I)—(III), and, moreover, in accordance with one of the formulae (1)—(4), as indicated by the subscripts. The transformation yields

$$p \supset q . = . p | q | q \qquad (5)$$

which, in view of our three rules of inference, is a *logical consequence* of (1)—(4). As (5) might have been chosen as a *definition* of the symbol " $\supset$ " in Sheffer's notation, we have exactly the same result as before : what can be derived from definitions (and logical truths), is a definition. If, however, we were to rest content with these examples and enunciate it as a general principle that whatever follows from a definition is a definition, we should be mistaken ; as can be seen from the example

$$2 = 1 + 1 \quad \text{Df} \qquad (1)$$
$$3 = 2 + 1 \quad \text{Df} \qquad (2)$$
$$4 = 3 + 1 \quad \text{Df} \qquad (3)$$
$$a+(b+1) = (a+b)+1 \quad \text{Df} \qquad (4) \text{ (recursive definition of addition)}$$

With the help of these four definitions it is easily proved that $4 = 2+2$:

$$4 =_{(3)} 3+1 =_{(2)} (2+1)+1 =_{(4)} 2+(1+1) =_{(1)} 2+2.$$

The result

$$4 = 2+2$$

is, admittedly, derived from mere definitions ; yet to regard it

as a *definition* of " 4 " would be most unnatural. Why not regard $8 \times 7 = 56$ as a definition of the number 56 ? Why, indeed, not regard *any* numerical equation as a definition ?

In this way we come to see that it is only in *certain* cases true to say that what follows from a definition is a definition, whereas in other cases it is not. Why this should be so is very puzzling, and we shall have to go into this point more fully. However that may be, the examples adduced are, I think, sufficient to dispose of the view that an analytic statement is one that *follows from the definitions* of its terms.

The true state of affairs can be expressed in a somewhat different way. Instead of saying "A statement is analytic if it *follows* from definitions ", we shall have to say "A statement is analytic if it can, by means of mere definitions, be *turned into a truth of logic*", i.e., if it is *transformable* into such a truth. Consider an example. Suppose I define a planet as a heavenly body moving round the sun. (This is not quite accurate as it fails to distinguish between planets on the one, and asteroids, comets, meteors, etc. on the other hand; and further because it ignores the fact that some other fixed stars also have planets. For the sake of simplicity let us, however, disregard these complications and keep to the simple definition). If I now say, "All planets move round the sun ", I am making an analytic statement. This statement is such that, in virtue of the definition, it can be *turned* into a logical truth. Indeed, replacing " planet " by its *definiens*, we get "All heavenly bodies which move round the sun move round the sun ", which is precisely the sort of truism one would expect to find in an analytic statement. Although we can see " with the naked eye " that this is a logical truth, we cannot yet identify the skeleton of the sentence in question with some definite logical formula, i.e., we cannot put a finger on the precise spot in PM, or in any other textbook of logic, saying, " *That's* the logical form of the statement." We shall rather have to transform our sentence by a number of steps until the logical skeleton of the proposition it expresses can be seen with perfect clarity. And in doing this, I hope to make an incidental gain—to throw more light on the status of a definition. Up till now the behaviour of a definition must have appeared rather erratic ; we shall understand the deeper reason for this far better when we come to see a definition in its natural setting of similar and related structures. To this end I shall make use of a somewhat new notation which will help to throw this point into relief.

Let p stand for the original sentence, "All planets move round the sun ". The first step is to translate the " all " idiom into

the idiom "there is no such thing that not"; for instance, "All men are mortal" can be paraphrased as "There is no man that is not mortal". Call this transformation T. Applying T to the sentence p in the way of an operator, we obtain

Tp = There is no planet that does not move round the sun. The next step is to put the term "planet" in the place of a predicate; let L symbolize this process, and write simply LTp for $L(Tp)$:

LTp = There is no thing such that it is a planet that does not move round the sun. Next eliminate the last restrictive clause and put it conjunctively (operation R):

$RLTp$ = There is no thing such that it is a planet and that it does not move round the sun. According to the principle of double negation (N), this can be expanded to saying:

$NRLTp$ = There is no thing such that it has not not the following property: it is a planet, and it does not move round the sun. Now translate the idiom "there is no such thing that not" back into the "all", or "whatever"—idiom (transformation T^{-1}):

$T^{-1}NRLTp$=Whatever a thing may be, it has not the following property: it is a planet, and it does not move round the sun.

According to the rule of De Morgan (M) this is further equivalent to

$MT^{-1}NRLTp$=Whatever a thing may be, it has the following property: it is not a planet, or it is not the case that it does not move round the sun.

Apply N^{-1} (the converse of N):

$N^{-1}MT^{-1}NRLTp$=Whatever a thing may be, it has the following property: it is not a planet, or it moves round the sun.

Now the denial of "If something is a planet, it moves round the sun" is "Something is a planet, and it does not move round the sun"; as the denial of the latter, according to M and N^{-1}, is "Something is not a planet, or it does move round the sun", and as the two denials cancel (N^{-1}), this must come to the same as the first sentence; which shows that the "if" idiom can be transformed into the "not" and "or" idiom (I):

$IN^{-1}MT^{-1}NRLTp$ = Whatever a thing may be, it has the following property: if it is a planet, it does move round the sun.

Let finally D stand for the definition of "planet"; applying D as a sort of operator, we get:

$DIN^{-1}MT^{-1}NRLTp$ = Whatever a thing may be it has the following property: if it is a heavenly body that moves round the sun, then it moves round the sun.

Repeating the R-step yields :

$RDIN^{-1}MT^{-1}NRLTp$=Whatever a thing may be, it has the following property : if it is a heavenly body and if it moves round the sun, then it moves round the sun.

Now at last we have reached the stage where the structure of our statement can be seen to coincide with a quite definite formula in PM, namely, with

$$(x) : \phi x . \psi x . \supset \psi x .$$

In words, whatever has the property ϕ and the property ψ, has the property ψ. This collects together a range of statements, each of which is of the form p. q. ⊃ . q, thus clearly exhibiting its tautologous character. In this way it is finally seen that the statement under consideration is a truth of logic. (In the transformation just carried out only the more important steps have been accounted for ; some of the purely idiomatic steps of paraphrasing have been neglected or telescoped.)[7]

It may seem pedantic, indeed excessively so, to go to such lengths to prove a very trivial thing. However, it helps us to see one point which we should not so easily have seen otherwise, viz. how similar the use of a definition is to that of other tools of transformation. Indeed, in looking back on the whole chain of transformations we have carried out, it becomes clear how near a definition comes to any of the other operators—as far as its *function* goes ; and how unnatural, for this reason, it would be to separate sharply the concept of a definition from that of other transformers.

A word must here be said on the notion of an *operator*. Suppose we consider any logical equivalence, i.e. an expression of the form

$$(. \ . \ . \ .) \equiv (. \ . \ . \ .)$$

As we pass along the equivalence, we are led from one expression to another without change of meaning ; now there are two directions of doing so, from left to right, or from right to left. Correspondingly, each logical equivalence gives rise to a pair of inverse operators, say Ω and Ω^{-1}, each of which can be used for transforming an expression into an equivalent one. Conversely, any such operator may be re-written as an equivalence, read from left to right, or from right to left, as the case may be. In this way operators and equivalences are

[7] Thus we have : there is a thing such that = there is something such that = there is at least one thing such that = there are some things such that = at least one thing exists such that = things (or : objects) exist such that, etc. Or again : there is no such thing that = it is not true (or : it is not the case) that there is a thing such that, etc.

intimately related. Thus we have the following correspondences:

N: $p \equiv \sim \sim p$
M: $\sim(p.q).\equiv.\sim p \vee \sim q$
I: $p \supset q.\equiv.\sim p \vee q$
T: $(x).\phi x.\equiv.\sim(\exists x).\sim\phi x$
etc.

I do not mean to say that an operator *is* an equivalence; an operator is, as observed, rather the *transition* from one expression to another in *accordance with* an equivalence. But as there is a one-one correspondence between pairs of operators on the one hand, and equivalences on the other, we shall in future not always take the trouble to distinguish between the two.

In formal respects it will be observed that two inverse operators, applied successively, yield identity, in symbols $AA^{-1} = 1$, and further that

$$(A\ B)^{-1}=B^{-1}\ A^{-1}.$$

However, it is not the subject of this paper to construct a calculus of operators.

A definition fits into the same scheme on the ground that it, too, can always be written as an equivalence; thus instead of defining the term "planet" in the usual way, we might have laid it down as an equivalence

x is a planet $\equiv$ x is a heavenly body moving round the sun.

Having said all this, we are now in the position to bring to light the fallacy that lies at the root of Schlick's and Ewing's definitions of the term "analytic". Two different concepts have been confused, namely, "to *follow from* a definition" and "to be logically true *in virtue* of a definition". The point has been made quite clear by W.V.O. Quine in an article published 1936.[8] "What is loosely called a logical consequence of definitions is therefore more exactly describable as a logical truth definitionally abbreviated: a statement which becomes a truth of logic when *definienda* are replaced by *definientia*". The same point, however, was already seen by Frege 1884 when he wrote.[9] "When a proposition is called . . . analytic in my sense, this is . . . a judgment about the ultimate ground upon which rests the justification for holding it to be true. This means that the question is . . . assigned, if the truth concerned is a mathematical one, to the sphere of mathematics. The problem thus becomes that of finding the proof of the proposition and of following it up

[8] "Truth by Convention": *Philosophical Essays for A. N. Whitehead*, p. 92.
[9] *Grundlagen der Arithmetik* trans. Austin § 3.

right back to the primitive truths. (*Urwahrheiten*). If, in carrying out this process, we come only on general logical laws and on definitions, then the truth is an analytic one, with the proviso that we must also take into account all propositions without which any of the definitions would become inadmissible. If, however, it is impossible to give the proof without making use of truths which are not of a general logical character . . . then the proposition is a synthetic one ".

2. Logical and idiomatic equivalence; definition and substitution licence

In this section we shall consider more closely how logical equivalences, definitions and other operators are related. We have seen how, by means of operators, a certain sentence can be transformed into a truth of logic. In retrospect two points stand out clearly :—(1) The sentence is transformable into such a truth not *only* by means of definitions, but with the material aid of certain other operators ; (2) the definition D behaves, in point of application, *not so differently* from operators such as N, M, I, T, L, R etc. These operators, however, fall into two distinct types which we will call *logical* and *linguistic*, respectively. Thus N, M, I, T are of the first, L and R of the second type. The principle according to which this distinction is made deserves perhaps some attention. The first group consists of those transitions which are valid on *logical* grounds alone. If they are written as equivalences, they become instances of certain logical truths. But the operators of the second group are not purely logical, and cannot be expressed in logical symbols only; rather it is characteristic of them that they are due to *the way word language is used*. To give an example for the first type, the operator N is the transition along the equivalence (indicated by the arrow).

$$\overrightarrow{p . \equiv \sim \sim p .}$$

and this equivalence is true in virtue of its *logical form ;* and similarly in the cases of M, I, and T. On the other hand, the operator L which enables us to pass from saying " There is a planet that moves round the sun " to " There is a thing such that it is a planet that moves round the sun ", putting the term " planet " in the place of a predicate, will be recognized by any user of the English language as idiomatically correct ; but a logical truth it is not. In fact, it is used to *prepare* the first sentence for symbolization within the frame work of symbolic logic, though it does not go far enough for that ; it's only

when we apply the further operator R (which eliminates the last restrictive clause), that we reach a sentence-form which can so be symbolized. From this it appears that the last two operators have a job very different from that of the former ones : they are transitions *within* word language, used, among other things, for re-phrasing a sentence such that it can go into symbols. The equivalence :

There is a planet that moves round the sun $\equiv$ There is a thing such that it is a planet that moves round the sun

is certainly true ; but it is true *neither on empirical nor on logical grounds ;* there is no formula in PM which covers such a case. It is true simply because, *according to the idiomatic use of the English language,* the two sentences come to the same. A *logical* equivalence (such as that linked with *M*) is *always* and *universally* true, irrespective of the language to which it is applied ; an equivalence of the latter sort, if it is true, is true because it is in accordance with the *particular way* in which a *particular language* (such as English) is used, but it holds no place in a universal system of logic. The two groups of equivalences are therefore not of the same standing. And yet *both* sorts are needed for transforming a certain sentence into a truth of logic. That shows that Frege's and Quine's account of the matter is incomplete in an important respect. According to Frege, only *general logical laws* and *definitions* are permitted in testing the analyticity of a proposition ; Quine stresses one thing only, *definitions*. Neither of them seems to have noticed the need for a third kind of processes which are *linguistic* in nature. If we were to limit ourselves to definitions, or to definitions and logical laws, we should never be able to translate the sentence "All planets move round the sun " into a truth of logic. Frege may have been led to the view he has taken by concentrating on mathematics to the exclusion of word language. A definition which is meant to apply to word language also, must obviously allow for other tools of transformation as well.

All this makes it desirable to modify our definition by saying : A statement is analytic if it can, by means of mere definitions, logical and, further, idiomatic (linguistic) operators, be turned into a truth of logic. We proceed now to consider these means separately.

The idea of a logical operator seems sufficiently clear insofar as it is based on that of a logical equivalence. A definition can always be re-written as an equivalence (see example above). However there is an important difference in that a definition is valid in virtue of the way a certain term, e.g., the word " planet",

is used in English, without, however, belonging to the body of truths provided once and for all by logic. In this respect a definition is more like an equivalence of the third kind. What is the difference ? If I say that the sentence, " Some planet moves round the sun " expresses the same fact as the sentence, " Something is a planet that moves round the sun ", I am making a statement about the way two phrases are used, viz., I am stating that the two sentences come to the same. This can, if we like, be construed as a *substitution licence* which gives permission to interchange these two locutions. The difference between a definition and a licence of the last sort is this : whereas a definition refers to *one* term only, and usually provides for its elimination from a context, an idiomatic licence applies to *whole sentences*, or clauses of such, or syntactical constructions, offering them as alternative modes of expression. That the dividing line is not an absolutely precise one, that the two sorts rather shade off into one another,—e.g., what B. Russell calls a "definition in use" also refers to a whole phrase —lies only in the nature of things and reflects how closely related, at bottom, the two sorts of licences are. Thus, whereas D is related to a *definitional* equivalence, and N, M, I, T to *logical* equivalences, L and R are related to *idiomatic* ones. By applying any of these three sorts of means, a statement, if it is analytic, is transformable into a logical truth.

In this way we have arrived at a definition, broad enough to be applicable to word language, and at the same time free from the defects which mar the afore-mentioned attempts. In the next section we shall have to consider some doubtful points to which the definition suggested gives rise. But before we do so it will be well go to into some minor points.

Let us consider the relation between a definition and a substitution licence more in detail, for the moment ignoring equivalences of the idiomatic kind. We shall approach this subject by asking first another sort of question : What is the difference between a substitution and a definition ? What, for instance, is the difference between $\frac{p \vee q}{\mathrm{p}}$ and $p \supset \mathrm{q}. = . \sim p \vee p$ Df ? The former is an instruction saying that $p \vee q$ is to be put in the place of p in the context of a *given formula only ;* it is *unidirectional,* i.e. it tells us to replace p by $p \vee q$, not $p \vee q$ by p, thus making that process *irreversible ;* and it *commands* us to carry out this substitution. The latter permits the substitution of $p \supset q$ by $\sim p \vee q$ in *any* context whatever, not just in that of one particular formula ; it allows us to pass *either* from $p \supset q$ to

$\sim p \vee q$, *or* from $\sim p \vee q$ to $p \supset q$, just as we please, making the transition *reversible;* and, finally, it *permits* it merely without actually instructing us to make it. We may express this briefly by saying that a definition is, not so much a rule (an instruction), as a *licence for re-writing* a sentence, or a formula, by putting *definiens* for *definiendum* and leaving it to us whether we wish to make use of it. In other words, a definition *paves a way* without forcing us to go it.

Similar remarks apply to *any logical* equivalence: we may pass from the one side of it to the other, in any direction and in any context we please; but we need not do so. Thus every logical equivalence supplies us with a substitution licence for replacing a certain expression by another one. In fact, all the operators which occur in our transformation chain, including the definition, provide us with such licences—further evidence of how closely related they all are.

This way of looking at things has, I submit, the advantage that it makes us see right from the start what matters in logic. It is not so much because of its *truth* that we take an interest in this or that particular equivalence, as because of the *use* we can make of it. What is really of value in a logical equivalence is that it lends itself to a use very similar to that of a definition: it supplies, to say it once more, a substitution licence, and that's in actual fact its more important side. But not *every* substitution licence supplies us with a *definition;* for a definition must be such as to permit the elimination of one symbol by others. Thus I can make use of the equivalence

$$p \supset q . \equiv . \sim p \vee q$$

in order to define the symbol "$\supset$" by putting,

$$p \supset q . = . \sim p \vee q \quad \text{Df}$$

If, however, I were to say

$$\sim \sim p . \equiv . p,$$

D. N.

I should, it is true, be stating a logical equivalence, but there would be no point in using it as a definition by writing

$$\sim \sim p . = p . \quad \text{Df}$$

For the latter formula may well serve to eliminate the combination of symbols "$\sim \sim$", but not the symbol "$\sim$" alone, and can, for this reason, not be taken as an explanation of that symbol. Similarly

$$\sim \sim \sim p . \equiv . \sim p,$$

though a perfectly valid equivalence and a substitution licence, would be useless as a definition of the symbol "$\sim$".

Shall we, then, demand of a definition that it should allow the complete elimination of one of the symbols? That won't do either. Suppose I define addition recursively by writing

$$a+(b+1) = (a+b)+1 \text{ Df},$$

then the symbol " + " occurs on both sides of the definition; yet it would be a mistake to reject, on this ground, the definition as circular. For the whole point of the formula is that it reduces the addition of a and $b+1$ to the simpler addition of a and b, and, if this process is repeated a suitable number of times, finally to the operation " +1 "; thus we can, step by step, reduce 7+5 to 7+4, 7+4 to 7+3, and so on, until we reach 7+1; the operation " +1 " itself is undefinable in arithmetic, being the simple step of forming the successor of a given number. We may bring this out, perhaps somewhat more clearly, by writing

$$a+S(b) = S(a+b) \text{ Df},$$

i.e., the sum of a and the successor b is the successor of $a+b$. This formula exhibits the general scheme according to which, in any particular case, the sum of two numbers can be defined in terms of " successor "; it is perhaps best characterized as an *instruction for framing particular definitions,* but is commonly itself taken for a definition by recursion. This example shows that a definition *need not* be a licence for eliminating a symbol. And even if a formula *does* permit eliminating a symbol, it need not be a *definition*. Take, for instance, the formula $4 = 2+2$; though it permits the elimination of the symbol " 4 ", it would hardly be recognized as a definition of this number. Thus the condition suggested is neither necessary nor sufficient. Whether conditions can be specified which determine, without any doubt, what a definition is, I do not know. Incidentally, in view of the indefiniteness with which nearly all the terms of word language are used, and the need for leaving, at least, some freedom for adjusting them to new situations which may crop up—would Aristotle have considered the case of a recursive definition?—I doubt the wisdom of pressing for a hard-and-fast rule, which can lead only to a sort of pseudo-precision. It is perhaps better to keep a term like definition flexible and make a decision, if the need arises, only in individual cases without anticipating the issue. That, of course, applies only to *natural* language; in an artificial, formalized language, the matter may be different.

We may sum up the discussion by saying that definitions are substitution licences of a *particular sort* (leaving the sense of this somewhat open), and that every substitution licence can

be re-written as an equivalence. It is only when expressed in this way that we can do logical work with definitions, for instance derive consequences from them.

We can now understand—what we failed to understand before—why it is that definitions behave in such a disorderly manner. The point is that a definition, if it is explicit, can be re-phrased as an equivalence. Now from an equivalence another equivalence may be derived by logical inference; if the latter is such that it can be used as an explanation of one of the symbols involved, it can itself be construed as a definition; thus it happens that the "logical consequence" of a definition may again be a definition. If, however, the equivalence obtained falls short of this demand, then what follows from a definition will not be a definition. In the former case the procedure by which we "deduce" a definition from another definition involves, strictly speaking, three separate steps: (1) putting the definition in the form of an equivalence, (2) deriving another equivalence from the given one by logical inference, and (3) re-writing the result obtained in (2) as a definition.

(*To be continued.*)

University of Oxford.

THE DESIRED AND THE DESIRABLE

By PEPITA HAEZRAHI

J. S. MILL'S dictum: "The sole evidence it is possible to produce that anything is desirable is that people actually desire it"[1] has been so often refuted on so many plausible grounds that an attempt to vindicate it might not be wholly devoid of interest.

One can recognise at least three distinctions of meaning between "desired" and "desirable" which either singly or in various combinations cover all ordinary senses of the two notions. These, I shall, for want of a more precise apellation, term (*a*) "the dispositional", (*b*) "the emotive or propagandist" and (*c*) "the imperative or absolute." I shall assume that "X is desired" means in all cases "Y is now desiring X", i.e. X has overcome all rival claimants and all obstacles to Y's fullest attention. "X is desirable" however seems cap-

[1] Utilitarianism, Ch. IV.

able of three different interpretations to match the three distinctions. Thus " X is desirable " may be equivalent to : (*a*) " X is capable of being desired " or " X would be desired if . . . " ; (*b*) " I approve of my (or your) desiring X " ; and (*c*) " X is good, therefore one ought to desire X ".

I wish to maintain that a consideration of the ways in which "desired" and "desirable" are actually used will show that far from being irreconcilable the difference between the two terms is one of degree rather than of kind and is adjustable to the point of interchangeability.[3]

(*a*) *The dispositional sense of " desirable "*

This is easily proved, indeed hardly a matter of dispute, in the context of the first interpretation. Here " X is desirable " i.e. " X is capable of arousing desire " strongly suggests the qualifying clause " but does not at this moment ". How, then can I know that X is capable of arousing desire ? Either by recalling that X has done so on some past occasion or occasions ; or by its resemblance to an object which is actually desired now or which has been actually desired in the past. Thus, though I usually prefer brandy I might refuse it on an evening dedicated to whisky ; i.e., though I usually find brandy desirable I now desire whisky. " The desirable " is equated with the " once, or usually desired ". It might even be actually present in the form of a very faint desire, accompanying or called forth by the recollection of a past desire and immediately overridden by the much stronger desire for my present drink or by the desire to be reasonable and not mix them.

This use of " desirable " is sharply brought out by examples like " It would be desirable for me to do some work but I'd rather go to the party ", or in a more famous formulation " Video meliora proboque deteriora sequor ". Here " X is desirable " means " X fits into my general system of values, I approve of X, indeed I do actually desire it ; but alas, I entertain

[3] It should be pointed out that such gradual adjustments take place where the " desired " and the " desirable " are considered against the concrete background of their actual relationships. The abstract logical relations between these two concepts remain however unaffected by such adjustments and, indeed, follow laws of their own. Thus in the context of the first interpretation—the most easily adjustable in its temporal concrete aspect, the logical relation is one of incompatibility and mutual exclusion : If X is desired (i.e. actualised) it is no longer desirable (potentially desired). In the context of the second interpretation a relation of one-sided dependence holds. The desirable is always something that is already desired. But the desired can be either desirable or undesirable. In the context of the third interpretation the desirable is completely independent and indifferent to the desired. For the desirable is defined as the good, and remains so regardless of whether it be actually desired or not. It is the object of this paper to show that this abstract relation, arrived at by a definition in vacuo, as it were, is not valid in the domain of actual relationships.

some fiercer desires which, being incompatible with X, induce me to give it up (good as it may be, fitting as it may be, desired as it may be) in favour of their own realisation ".

It should be pointed out that the reference of "desirable" in the above examples to fittingness or approval is *fortuitous* and *by no means necessary.* "Cigarettes are desirable but I desire health more and therefore will not smoke them" and "Health is desirable but I desire a smoke more and will have my cigarette" are two equally justifiable assertions. Asuming health to be indeed the greater good, this example neatly brings out my point: "desirable", in this context, does not necessarily refer to a greater good; it connotes solely an over-ridden and supplanted desire. To misquote a saying of Gregory of Nazanz' which the Master of Balliol is very fond of citing:[3] "The actually desired is not necessarily the undesirable".

The difference, then, between "desirable" and "desired" lies in the time of realisation, or in intensity and degree. The ultimate reference of the desirable (indeed its sole evidence for being desirable) is *either* to a past or future actual desire which X has or expects to be aroused *or* to a present actual, but very faint and overridden desire.

(*b*) *The emotive or propagandist sense of "desirable"*.

"Desirable" is often employed in a hortatory manner.[4] At such times use is made of the double-entendre of this term with its simultaneous reference to desire and to approval, to stress the latter and thereby increase its emotive and propagandist power. An additional sleight of hand is performed by presenting this approval as applying directly to X, when in fact it applies only to the "desiring of X". For "X is desirable" in this context means either "I approve of my desiring X" or "I approve of your desiring X". It is of great interest to note that only the first case implies a *necessary* approval of X; the second tacitly assuming it on insufficient grounds.

Thus, when, being an addict, I say "Music is desirable", what I mean is: "Music is a good thing; I enjoy and desire it very much; I think well of myself for desiring it, I hope I will go on desiring it". On the other hand when disliking a certain task and wishing someone else to undertake it one says, *e.g.*: "Domestic service is desirable", one appears to convey a

[3] "The kingdom of heaven is not necessarily confined to fools."

[4] I deal here with the hortatory use of "desirable" only, i.e. cases where existing desires are checked or encouraged. I shall deal with the imperative sense of desirable which claims to order non-existent desires into being later. Obviously a strong hortatory flavour adheres to the imperative use; but not vice versa.

favourable attitude to domestic service but what one has is a favourable attitude to the desiring of domestic service by someone else. One says, in effect, " Domestic chores are a *bore*, how wonderful if someone else would do them ".

This is a rather extreme example of the demagogic misuse to which " X is desirable " lends itself. Milder examples are met with daily as one says : " Early bed-times are desirable ", or " Regular attendance at Church is desirable ", when one has no intention of complying oneself but, on the other hand, has no actual distaste for these activities.

Altogether four uses of "desirable" in the hortatory sense can be distinguished : (*a*) " I desire X and desire to go on desiring X " ; (*b*) " I desire X and desire somebody or everybody else to desire X " ; (*c*) " Though I do not desire X I desire to desire it ;[5] " (*d*) " Though I do not desire X myself I desire somebody else to desire it ".

In all four cases the actual object of desire and approval is the " desiring of X " not X. This is usually misunderstood and desire and approval are assumed to apply directly to X. Sometimes, X *happens* to be *also* approved and desired and then this hasty assumption is rendered innocuous. But sometimes X is not desired and approved, and if the spurious effects of its being desired and approved be consciously misused, the results might prove pernicious.

But the sole point which need concern us here is that " X is desirable " in the hortatory sense refers to, and is directly dependent on an existing actual desire for the "desiring of X", which it tries to intensify and perpetuate (or check and lessen when " undesirable " is used).

(*c*) *The imperative or absolute sense of desirable.*

It is the third sense which is commonly supposed to exemplify the irreconciliablity of the "desired" and the "desirable ". "X is desirable" is equated to "X is good, X ought to be desired regardless of whether it is actually desired or not ". In the second sense, when X is not actually desired, the imperative or absolute interpretation retains its full stringency. Thus, " X is desirable " *i.e.* " X ought to be desired " seems to imply an obligation to desire a certain object though it had never been desired before, is not desired now, and is quite incapable of

[5] This is the hortatory equivalent of " video meliora ", if a very faint desire for X is existent which I desire to render sovereign ; or the hortatory equivalent of the imperative or absolute interpretation if I try to call forth a non-existent desire, by my desire for this desire. But of this later.

arousing desire even under the most auspicious of circumstances. On asking why on earth we should be obliged to desire that object at all, we are told because it is " good " (or conducive to " good ")[6] and that one should always desire what is good (or conducive to good). In short, " desirable " is the predicate of an X which, though a fit object of a judgment of approval, *i.e.* " good ", *does not* affect our conative tendencies at all. By introducing an alien principle of valuation, the norm of the " good ", an absolute difference in kind is assumed between the desired and the desirable.

To illustrate this point let us imagine something " good " which possesses no reference to desire; *e.g.* a certain piece of modern music. We approve of the composition in so far as it satisfies the criteria, rules and principles of what good music ought to be. Now let us assume that listening to it affords none of the pleasures associated with classical music and leaves us no desire to repeat the experience, and also fails to satisfy any desire for self-improvement or for novel experience. In short, the composition is " good " because it comes up to standard, but we do not actually desire to listen to it for any reason whatsoever. This desirable—because good—piece of music would then not be desired at all and at first blush the case would seem proved for a use of " desirable " and " desired " in the third sense.

Now this argument has a very serious flaw: it fails to account for the difference of meaning between " good " and " desirable ". The hortatory power which judgments of approval (*i.e.* " good ") possess in themselves is, for instance, considerably intensified[7] and to a certain degree actualised and activated by the epithet " desirable ". When instead of calling X " good " I call X " desirable " I wish to imply in addition to: (*a*) that X is good; (*b*) that the good ought to be desired (which are contained in " good "); (*c*) that the good *can* be desired: other good things in my experience have been desired; at this very moment, I desire a good thing and desire it to be desired by others;[8] (*d*) that in a very small way it is in my power to bring a state about in which this thing will be actually desired by others, *i.e.*

[6] Both alternatives equally satisfy the argument.

[7] It is interesting to note that in this case the hortatory intensification draws its power from a reference to the actually " desired " and not as in the case discussed above from a reference to approval. This sheds new light on the inherent ambivalence of " desirable " (i.e. its simultaneous reference to " good " and to " the desired ") and on the possibility of a demagogic use of this double-entendre, as at the dictate of interest and not always in accordance with one's real intentions, now the one, now the other aspect is thrown into relief.

[8] But of this later.

by calling it "desirable" I recommend it as if it were already and generally desired; (*e*) that I actually, at this moment intensely desire to bring such a state about. The "desirable" in the imperative sense of "what ought to be desired" includes therefore a strong and essential reference to what is actually desired. That one ought to desire X, means that somebody exhorts one to desire X, in other words that somebody who actually desires X himself also desires others to desire X. Even the rejected piece of music must presumably have been desired, and desired passionately, by at least one person—its composer.

I cannot at this moment recall a single example of an object, which, having been considered desirable (and good), has not actually been desired by at least one person: the person who discovered it to be so. Thus, a social reformer who propagates the idea of absolute equality, desires people to desire absolute equality. He thinks that they ought to desire it. He himself approves of it and actually desires it. And though all men or the majority of men do not as yet desire it in actual fact, he does his best to bring about such a state.

The ultimate reference of the "desirable" in the sense of "what ought to be desired" is therefore to the actually desired, namely, the actually desired by the person who dictates the "ought" and brings into prominence the good hitherto neglected, overlooked or rejected. The difference between "desired" and "desirable" appears once again not so much a difference in kind, as in the generality of distribution; in short, a quantitative difference adjustable by degrees.[9]

This seems to me to prove conclusively that "the sole *evidence* for something being desirable is that *people* actually desire it."[10] We must be careful though to point out that *evidence* does not mean *reason*. We know that X is desirable because we find ourselves: (*a*) desiring X and encouraged in this desire; or, (*b*) being admonished to desire X by somebody who having found it good, actually desires X; or, (*c*) being admonished to desire it by somebody who does not think X good, does not desire X but misuses the hortatory power of the word to mislead us as to the nature of X and of his feelings about X. But this is not the reason *why* X is good, nor why it is desirable. It is but the proof *that* X is desirable, *i.e.* of an existing relation between X and desire. The reasons are the actual qualities of X,

[9] Here, as in the dispositional use of the term, what is desirable is what has not been fully realised. But whereas there the full realisation inside *one single person* had been impeded by circumstantial factors or lack of intensity, here its chief lack is scope; i.e. though fully realised (intensely desired by one person) it is not generally accepted and desired by others.

[10] The italics are mine.

which do not concern us here. This may possibly explain why the predicates " desirable " and " desired " are not fool-proof signs for the axiological standing of an X: " the good is not necessarily the undesired "(!)

The weakness of the argument is in the indeterminate denotative scope of the term "*people*". If this be reducible to its lower limit of one, then Mill's dictum holds. But if it be taken to mean "a reasonable majority"[11] then it must be admitted a moot point.

In other words as long as we take Mill's dictum to be a statement about the desirable being good by virtue of the third interpretation; and the evidence for the good being desirable its being actually desired—it must be granted that this is the only way in which the desirable (and the good) can be recognised at all. But the claim that what the majority (by the way which ?)[12] desires is therefore desirable *and* good, is to say the least highly questionable. Therefore we ought to be very careful to keep these two assertions apart and be fully aware whether we argue against the first or the second. This seems to me a point of some importance, when considering such arguments as are brought forward by C. L. Stevenson in *Ethics and Language*. He writes:—

" . . . J. S. Mill . . . treats the statement " *If something is desired it is desirable* ", as though it were axiomatic, the antecedent being the " sole evidence " it is possible to produce for the consequent. If " desirable " meant capable of being desired, the statement would indeed be innocent enough. But Mill intends the word to carry all the import of " good ". Thus understood the statement so far from being axiomatic becomes highly controversial . . . " *That which is desired is desirable* " is a statement characteristic of an easy-going man who wishes to encourage people to leave their present desires unchanged and conversely the statement " That which is desired is *not* desirable" is characteristic of the stern reformer who seeks to alter or inhibit existing desires. *Statements about what is desirable unlike those about what is desired, serve not to describe attitudes* MERELY *but to intensify or alter them.*[13] The alleged axiom then is controversial because it leads to disagreement in attitude. Although it seems innocuous to those content with a ready status quo it is intolerable to those who are striving to make fundamental changes in men's aims; and it is particularly invidious because its concealed

[11] As no doubt Mill intended it to.
[12] The majority of which country, which continent, which world ?
[13] All italics are mine.

pun makes it seem to give the former people alone an axiomatic support. In defence of Mill it must be mentioned that he was assuredly not seeking this effect . . . ."[14]

My conclusion that statements about the desirable no less than those about the desired describe actual attitudes (namely, those of reformers and innovators of all kinds) seems born out by Stevenson's choice of the qualifying adverb *merely* in the italicised proposition (*vide supra*). That is, Stevenson holds without reservation that " if something is desirable it is also desired ". Now this is really all that Mill claims on the first point, though he also seems to imply by using the term "people" that "what is desirable is *generally* desired." Nowhere does Mill assert what Stevenson by a very elementary error in logic[15] takes him to assert; namely, that " if something is desired it is desirable ". This Stevenson characterises as the attitude of the easy-going man and proceeds to attack. But surely no man is that easy-going. Surely some of the things actually desired must appear undesirable even to the easiest-going (for instance the desire of reformers for reforms !) What singles out the easy-going man is that he is apt to consider the usual and general desires of normal people to equate pretty nearly with what he holds desirable. He is also inclined to take the actual and general presence of a desire as a point in its favour. In this he concurs with the distributive implication of Mill's dictum so that when Stevenson attacks his stand he argues against the second (distributive) not the (essential) first aspect of Mill's dictum. The first essential aspect Stevenson does not touch upon except, unconsciously, and in a somewhat off-hand manner to agree with it.

To conclude : Mill's dictim is immune to attacks directed against his definition of the "desirable" (as implying the notion of "good" and "superior value") for in this context it seems impossible to find evidence that any desirable thing has been known to exist which has not also been actually desired by at least one person.

But the explicit inclusion of a reference to the distributive generality of being desired shifts the point of the argument. Mill goes on to argue that happiness is desirable not because it is good and because somebody desires it, but because *each* person and therefore *all* persons desire it. The proof for the desirable (what Mill desires us to desire) being desirable (good)

[14] Ch. 1, p. 17.

[15] (*a*) (If something is desirable it is desired) entails (*b*) (All that is desirable is desired). But (*b*) does not *necessarily* entail (*c*) (All that is desired is desirable) but only (*d*) (Some things which are desired are desirable). (*a*) therefore does not entail (*e*) (If something is desired it is desirable).

is not that it is good and that Mill desires it, but that *all* people desire it. Mill may not have sought this effect any more than the one of which Stevenson acquits him, but these are the ultimate elements to which his argument reduces itself, and the one form in which it is open to legitimate attack.

Bedford College, London.